Appendix B: Mean and Standard Deviation Calculations Using a Scientific Calculator or a Computer

Statistical manipulation is often necessary to order, define and/or organize raw data. A full analysis of statistics is beyond the scope of this work, but there are some standard analyses that anyone working in a biology laboratory should be aware of, and know how to perform.

MEAN

The mean is the average of all the values obtained. It is computed by the sum of all of the values (Σx_i), divided by the number of measurements (n). If the sum of all numbers is 31, while there were 9measruement, the mean is 3.44.

$$\bar{X} = \frac{\Sigma x_i}{n}$$

When random data is recorded, a plot of frequency vs. occurrence of each measured value will result in a normal distribution curve . That is to say that for any random variance of a set of data, the data will be distributed in a bell shaped curve, where the mean is at the center of the curve and the most frequently-occurring measurement. The standard deviation of the mean is the measurement of how wide this bell shaped curve is. When the standard deviation is small, the data occurs within a narrow range of the mean value. When the standard deviation is large, the data reflects a widely varying range of measurements.

The standard deviation is given mathematically by the formula:

Standard Deviation = $\sqrt{\frac{\sum(x_i - x_{avg})^2}{n-1}}$ is the standard deviation (σ_{n-1})

where x_I represents the individual measurements
x_{avg} represents the mean
n represents the number of measurements

I. Statistical Analysis Using the TI-36X Calculator

A. PRELIMINARY

If the data consist of a **single** list of numbers, then list them in a column or in some other organized way so they can be entered systematically and with a minimum chance of error.

If the data consist of **paired** numbers, then list the numbers in a table consisting of two columns. List values of an independent variable in the first column and values of the other variable in the second column. The two paired values are always shown on the same row of the table.

To initiate an action identified by an instruction written in white directly on a calculator key (written on the lower half of the key), press that key directly.

To initiate an action identified by an instruction written in yellow directly on a calculator key (written on the upper half of the key), first press the **[2nd]** key, then press the key identified by the yellow label.

To initiate an action identified by an instruction written immediately above a calculator key, first press the **[3rd]** key, then press the key identified by the label above it.

B. GENERAL PROCEDURE FOR STATISTICAL ANALYSIS OF DATA

1. Turn on the calculator and clear storage registers as follows:

Press **[AC/ON]**. This clears the display and deletes numbers stored in various memory locations. In the solar model, the display turns on automatically when you expose the solar panel to light. However, the memory registers may contain numbers introduced during turn-on, so **[AC/ON]** should be pressed in any event. The calculator turns off automatically within 20 sec when the cover is placed over the calculator or if its top surface is otherwise darkened, clearing all memories and storage registers.

The **[CE/C]** key clears incorrect entries, error conditions and pending operations. However, this key does not affect information stored in memory registers. Press **[CE/C]** once to clear an incorrect numerical entry or an Error condition as indicated on the display. Press **[CE/C]** twice to clear the display and all pending operations.

2. Enter data as follows:

a. Sequence of operations if a single list of numbers is to be entered. (Refer to section 2 if numbers are to be numbered in pairs.)

Press **[3rd] [STAT 1]**. The word "STAT" should appear in the window.

Press number keys (white labels on gray keys) to enter the first number from the set of data.

Press [∑+]. This enters the displayed number into a storage register and displays the number "1", indicating that 1 unit of data has been stored in the statistical registers of the calculator.

Press number keys to enter the second number from the set of data.

Press [∑+]. The number "2" is displayed to indicate that 2 units of data have been stored.

Repeat this sequence of entering a number, then storing it with the [∑+] key, until all numbers have been entered.

After the last number of the data has been stored, be sure that the integer displayed by the calculator is the same as the total number of items of data entered into the calculator. If not, press **[2nd] [CSR]** to clear the calculator registers, then re-enter the data.

b. Sequences of operations if a list of paired numbers are to be entered. (Refer to section 1 if a single list of numbers is to be entered.)

Press **[3rd] [STAT2]**. The word "STAT" should appear in the display.

Press number keys (white labels on gray keys) to enter the value of the independent variable (x-value) of the first pair of number in the table of data.

Press **[X<->Y]**. This key stores the value of the independent variable temporarily and prepares the calculator to receive the value of the other variable.

Press the number keys to enter the value of the other variable (y-value) of the first pair of numbers.

Press **[∑+]**. This stores the two numbers as a unit of data consisting of an ordered pair and causes "1" to be displayed, indicating that one pair of numbers has been stored so far.

Enter the second pair of numbers by the same sequence of steps:

- enter the value of an independent variable, using number keys
- press **[X<->Y]**
- enter the value of the other variable, using number keys
- press **[∑+]**. The number "2" is displayed by the calculator.

Repeat this same sequence until all of the ordered pairs of numbers have been entered.

After the last pair of numbers has been entered, be sure that the integer displayed by the calculator is the same as the total number of pairs of numbers entered into the calculator. If not, press **[2nd] [CSR]** to clear the calculator registers and re-enter the data.

3. Perform statistical operation on the data entered into the calculator:
See pages A-3 and A-4 for some specific statistical operations.

4. Clear registers and turn off calculator as follows:
Press [AC/ON]. This removes all the data from the calculator's registers.

Replacing the cover will turn off the calculator by covering the solar panel. All storage registers will be cleared within 20 sec.

C. MEAN AND STANDARD DEVIATION OF A SINGLE SET OF NUMBERS

1. Calculating the MEAN of a set of numbers previously entered into the calculator.
The calculator must be in the "STAT" mode (indicated in the display) and the data to be analyzed must have been entered appropriately into the storage registers. (Instructions for entering the data are on pages A-1 through A-3)

Press **[2nd] [$\bar{x}$]**. This displays the mean of the previously entered numbers.

2. Calculating the STANDARD DEVIATION of a set of numbers previously entered into the calculator.
The calculator must be in the "STAT" mode (indicated in the display) and the data to be analyzed must have been entered appropriately into the storage registers. (Instructions for entering the data are on pages A-1 through A-3)

Press **[2nd] [$\sigma_{x\ n-1}$]****. This displays the mean of the previously entered numbers.

** The **[$\sigma_{x\ n-1}$]** key is used if the data you entered are a sample from a much larger set of numbers, any of which you could have selected for the data set. For example, if you are measuring the diameters of a few human red blood cells this represents only a few of all the normal human red blood cells in the world,

so the **[$\sigma_{x\ n-1}$]** function would be used. If you are analyzing a complete set of all items which could be used in the data set and not just a sample of them, then use the **[$\sigma_{x\ n}$]** key to determine the standard deviation. Foe example, if you are listing the standard deviation of the weights of all of the (85 or so at last count) whooping cranes in the world, then you would use the **[$\sigma_{x\ n}$]** function.

D. LINEAR REGRESSION ANALYSIS OF A SET OF PAIRED NUMBERS

1. Calculating the LINEAR CORRELATION COEFFICIENT (coefficient of linear correlation) of a set of paired numbers previously entered into the calculator.

The calculator must be in the "STAT" mode (indicated in the display) and the data to be analyzed must have been entered appropriately into the storage registers. (Instructions for entering the data are on pages A-1 through A-3)

Press **[3rd] [Cor]**. This displays the linear correlation coefficient of the previously entered set of numbers.

2. Calculating the SLOPE and Y-INTERCEPT of the regression line generated from a set of paired numbers previously entered into the calculator.

The calculator must be in the "STAT" mode (indicated in the display) and the data to be analyzed must have been entered appropriately into the storage registers. (Instructions for entering the data are on pages A-1 through A-3)

Press **[2nd] [SLP]**. This displays the slope of the linear regression line.
Press **[2nd] [ITC]**. This displays the y-intercept of the linear regression line.

These two values are the **m** and **b** of the equation of a line **y=mx+b**. You can use these two values in a variety of ways. You can use them to make predictions about unknown (or as yet unmeasured values of **y** or **x**) or you can use them draw the relationship between the two values (calculate a y for x values at the extremes of your data).

3. Calculating a Y-VALUE that would correspond to a specific x-value on the linear regression line generated from a set of paired numbers previously entered into the calculator.

The calculator must be in the "STAT" mode (indicated in the display) and the data to be analyzed must have been entered appropriately into the storage registers. (Instructions for entering the data are on pages A-1 through A-3)

Press number keys to enter any x value into the calculator.

Press **[2nd] [y']**. This displays the y-value of the point on the regression line that corresponds to the entered x-value.

You have asked the calculator to substitute a particular x-value into your **y=mx+b** equation and determine what the y-value is for that particular value of x.

4. Calculating a X-VALUE that would correspond to a specific y-value on the regression line generated from a set of paired numbers previously entered into the calculator.

The calculator must be in the "STAT" mode (indicated in the display) and the data to be analyzed must have been entered appropriately into the storage registers. (Instructions for entering the data are on pages A-1 through A-3)

Press number keys to enter any y value into the calculator.

Press **[2nd] [x']**. This displays the x-value of the point on the regression line that corresponds to the entered y-value.

You have asked the calculator to substitute a particular y-value into your **y=mx+b** equation and determine what the x-value is for that particular value of y.

II. SPREADSHEET ANALYSIS: Microsoft Excel

A. DATA ENTRY

This program is available on the computers in class (if you have computers in your classroom) or on other computers on campus. Data is entered in rows and columns. In Excel the rows are numbered and the columns are lettered. Data entry is performed by highlighting the cell (for instance the first cell is A1 in the upper left corner) and entering the value. If you have paired sets of data (such as concentration and absorbance) these values are placed in rows where each column is a kind of information. So all absorbance values might go in Column B, paired with appropriate concentration values in Column A. It is easiest if you place the independent variable in Column A and the dependant variable in Column B.

B. MEAN AND STANDARD DEVIATION

1. **Sum of a set of data.** Highlight the cell where you wish to the sum to appear in and click on "Insert" on the tool bar. Drag down the mouse to select "f_* Function". From the left-hand list that comes up, select either the "Most Recently Used" or "Statistics" function category. From the list that comes up on the right side, select "SUM" from the function names, and click on "OK". A screen will come up in which you must input the numbers that you wish to sum. Note: the colon mark ":" reads as a "to" in Microsoft Excel, so if you want to sum a whole row, you can name it by typing in the cells at the beginning and end of the row to be summed, separated by the colon mark. For example, cell A1 to A6 would read "A1:A6". Finally, click on "OK", and the sum should appear in the cell you selected for it. Alternatively, a short cut would be to simply select the cell where the sum is to appear, then type "@sum( )", where in the parentheses, the cells to be summed appear.

2. **Average (Mean) of a set of data.** This operation is nearly the same as that described above for obtaining a sum. Highlight the cell where you wish to the mean to appear in and click on "Insert" on the tool bar. Drag down the mouse to select "f_* Function". From the left-hand list that comes up, select either the "Most Recently Used" or "Statistics" function category. From the list that comes up on the right side, select "AVERAGE" from the function names, and click on "OK". A screen will come up in which you must input the numbers that you wish to sum. Input your cells to be averaged, and click on "OK".

3. **Standard deviation of the mean.** This operation will also work nearly the same as that described for calculating a sum and mean. Highlight the cell where you wish to the mean to appear in and click on "Insert" on the tool bar. Drag down the mouse to select "f_* Function". From the left-hand list that comes up, select either the "Most Recently Used" or "Statistics" function category. From the list that comes up on the right side, select "STDEV" from the function names, and click on "OK". A screen will come up in which you must input the numbers that you wish to sum. Input your cells to be averaged, and click on "OK".